

NASA Contractor Report 4451

IN-53
110800

P.32

A Principled Approach to the Measurement of Situation Awareness in Commercial Aviation

Yvette J. Tenney, Marilyn Jager Adams,
Richard W. Pew, A. W. F. Huggins,
and William H. Rogers

CONTRACT NAS1-18788
JULY 1992

(NASA-CR-4451) A PRINCIPLED
APPROACH TO THE MEASUREMENT OF
SITUATION AWARENESS IN COMMERCIAL
AVIATION (8BN) 32 p

N92-30306

Unclassified

H1/53 0110800





NASA Contractor Report 4451

A Principled Approach to the Measurement of Situation Awareness in Commercial Aviation

Yvette J. Tenney, Marilyn Jager Adams,
Richard W. Pew, A. W. F. Huggins,
and William H. Rogers
BBN Systems and Technologies
Cambridge, Massachusetts

Prepared for
Langley Research Center
under Contract NAS1-18788



National Aeronautics and
Space Administration

Office of Management

Scientific and Technical
Information Program

1992

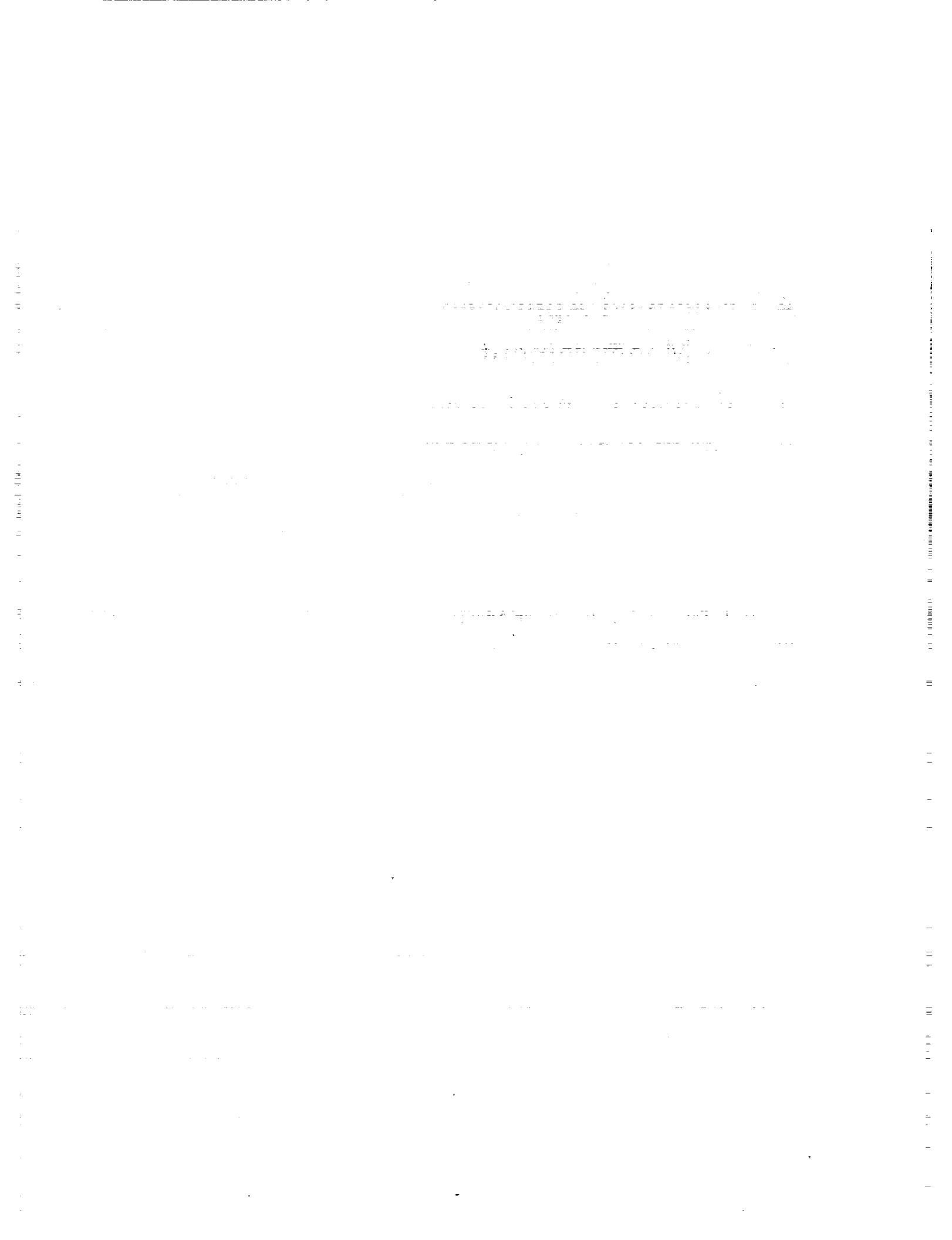


Table of Contents

| | Page |
|--|-----------|
| INTRODUCTION..... | 1 |
| Why Measure Situation Awareness?..... | 1 |
| Situation Awareness and Pilot Error..... | 1 |
| How Has Situation Awareness Been Defined?..... | 2 |
| A COGNITIVE VIEW OF SITUATION AWARENESS..... | 2 |
| Neisser's Theory of the Perceptual Cycle..... | 3 |
| Real-World Constraints on Situation Awareness..... | 5 |
| The Pilot's Job..... | 5 |
| The Available Information..... | 6 |
| Depth of Analysis Required..... | 6 |
| Change in Focus or Context | 7 |
| Reconceiving the Demands on the Pilot..... | 7 |
| Sanford & Garrod's Theory of Memory & Attention..... | 7 |
| Mission Categories..... | 9 |
| An Airline Accident | 9 |
| The Nature of Expertise | 10 |
| Shedding, Delaying, and Preloading Tasks..... | 11 |
| Letting Go of Unnecessary High-workload Strategies | 12 |
| Utilizing Non-competing Resources | 12 |
| Multiplexing Transmissions Over Time | 12 |
| Shortening Transmissions..... | 12 |
| Converting Interference from Concurrent Transmissions..... | 12 |
| Chunking Transmissions..... | 12 |
| Summary..... | 12 |
| STRATEGIES FOR THE MEASUREMENT OF SITUATION AWARENESS..... | 13 |
| Measures of Situation Awareness Derived from Scenario Manipulations..... | 13 |
| Design Scenarios with Subtle or Misleading Indicators | 15 |
| Design Scenarios with Interruptions | 15 |
| Direct Measures of Situation Awareness | 16 |
| Queries | 16 |
| Recall of Displays | 16 |
| Category Judgments | 16 |
| Information Seeking | 17 |
| Model-Based Measures of Situation Awareness | 17 |

INTEENTIONALLY BLANK

| | |
|--|----|
| ASSESSING THE IMPACT OF NEW TECHNOLOGIES ON SITUATION AWARENESS: TWO EXAMPLES | 18 |
| Example 1: Evaluation of Panoramic Flight Path Displays..... | 18 |
| Introduce Subtle or Misleading Information | 19 |
| Introduce Interruptions | 19 |
| Queries | 19 |
| Example 2: Fault-Finder..... | 19 |
| Introduce Subtle or Misleading Information | 20 |
| Introduce Interruptions | 20 |
| Queries | 20 |
| Recall of Displays..... | 20 |
| Category Judgments | 20 |
| Information Seeking or Recording of Eye Movements | 20 |
| AUTOMATION AND SITUATION AWARENESS..... | 20 |
| CREW AWARENESS | 22 |
| SUMMARY AND CONCLUSIONS..... | 22 |
| REFERENCES..... | 23 |

List of Figures

| | |
|---|----|
| Figure 1: Neisser's (1976) view of the perceptual cycle..... | 4 |
| Figure 2: Neisser's (1976) expanded view of the perceptual cycle..... | 5 |
| Figure 3: Modification of Neisser's (1976) view of the perceptual cycle..... | 8 |
| Figure 4: Hypothetical view of the pilot during landing. | 10 |
| Figure 5: Hypothetical view of the pilot during the incident described in the text | 11 |

List of Tables

| | |
|---|----|
| Table 1: Characteristics of Pilot Situation Awareness..... | 14 |
|---|----|

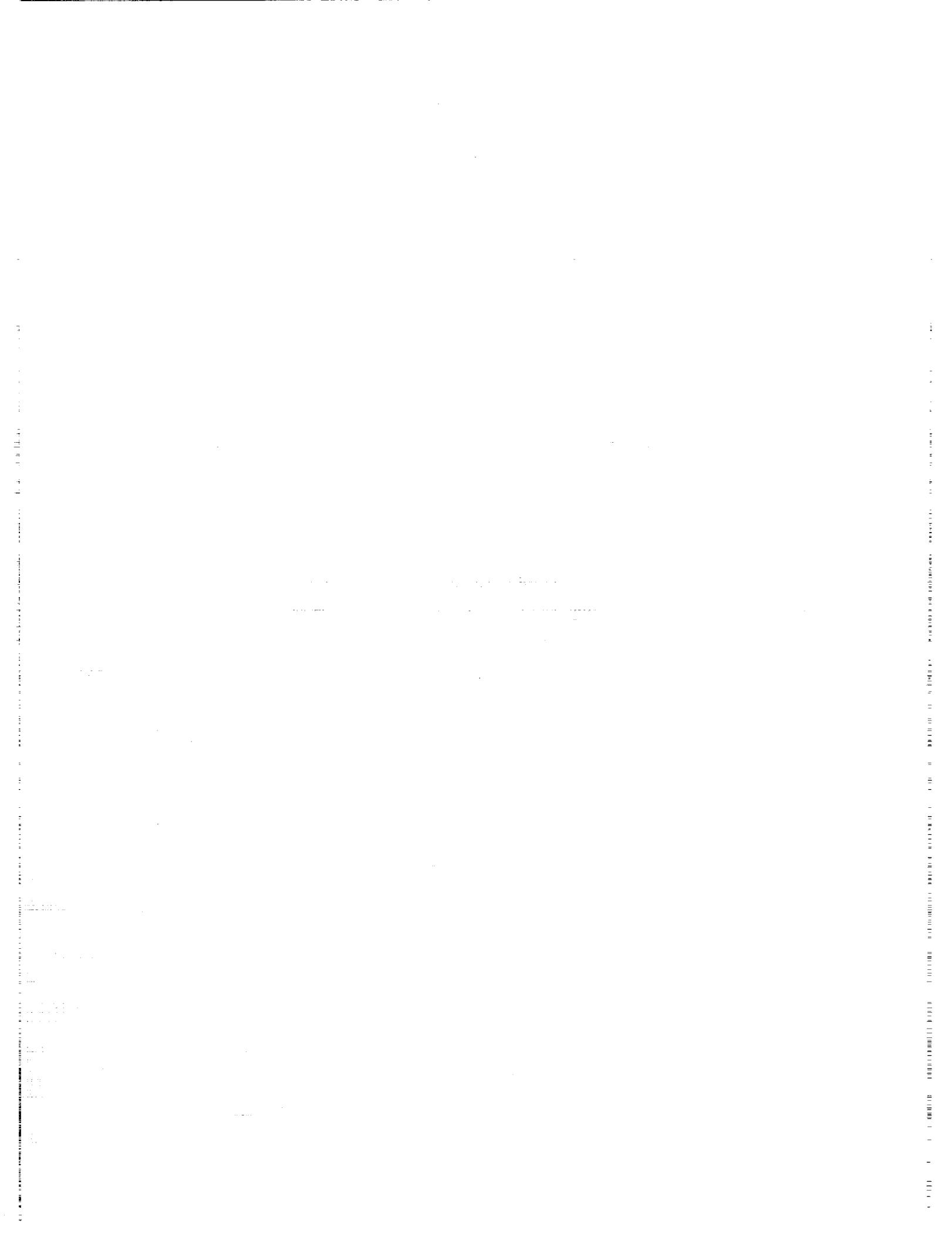
PREFACE

This work was performed under NASA Contract No: NAS1-18788 to Bolt Beranek & Newman, Inc.

We wish to thank Dr. Randall L. Harris, the NASA Langley Technical Monitor of this project, and Dr. Kathy Abbott, the NASA Langley Contracting Officer's Technical Representative, for their congenial oversight of the project. We also appreciate the technical discussions and contributions of Dr. Sheldon Baron of BBN and Mr. Russell Parish and Dr. Dean Nold of NASA Langley Research Center.

ABSTRACT

The issue of how to support situation awareness among crews of modern commercial aircraft is becoming especially important with the introduction of automation in the form of sophisticated flight management computers and expert systems designed to assist the crew. In this paper, cognitive theories are discussed that have relevance for the definition and measurement of situation awareness. These theories suggest that comprehension of the flow of events is an active process that is limited by the modularity of attention and memory constraints, but can be enhanced by expert knowledge and strategies. Three implications of this perspective for assessing and improving situation awareness are considered: (1) scenario variations are proposed that tax awareness by placing demands on attention; (2) experimental tasks and probes are described for assessing the cognitive processes that underlie situation awareness; and (3) the use of computer-based human performance models to augment the measures of situation awareness derived from performance data is explored. Finally, two potential example applications of the proposed assessment techniques are described, one concerning spatial awareness using wide field of view displays and the other emphasizing fault management in aircraft systems.



INTRODUCTION

The term *situation awareness* is increasingly finding its way into the literature on the human factors of system development. In everyday parlance, it means the up-to-the-minute cognizance required to operate or maintain a system. It is generally agreed that expert pilots maintain intimate and meticulously updated knowledge of the state of their aircraft and its support systems. Although such situation awareness contributes to good performance, it is not synonymous with it. It is possible to have good situation awareness and still not be a good pilot because of poor motor skills, coordination or attitude problems. Conversely, under automated flight conditions it is possible to have good performance with minimal situation awareness. Indeed, much of this knowledge may be superfluous in the case of routine flights. However, when emergencies arise and aids can no longer be relied upon, the completeness and currency of the pilot's situation awareness are critical to the ability to make decisions, revise plans, and manage the aircraft.

In aviation communities there is increasing concern about whether the appropriate level of situation awareness is being acquired and maintained by crews. One focus of attention is on the extent to which automation that is designed to reduce crew workload serves inadvertently to decrease their situation awareness. It is argued that as automation takes over functions previously assigned to the crewmembers, they lose contact with the information needed to perform and monitor these functions, the very information that is essential to effective decision-making in the event of a malfunction or emergency situation.

On the other hand, the advent of high technology or "glass cockpits" has opened up the design space to include "smart" information retrieval, large screen pictorial displays and even three-dimensional displays. These advances offer the potential to make it easier for

the crew to acquire and maintain certain aspects of situation awareness.

WHY MEASURE SITUATION AWARENESS?

Reason (1990), in his discussion of human error, distinguishes between "unsafe acts" and "precursors of unsafe acts." Our goal in developing techniques for measuring situation awareness is to help uncover information-processing precursors of unsafe acts. Analyses of scenarios that tax situational awareness may reveal shortcomings of user-aids that would not be evident in a less demanding scenario. Similarly, the use of probes and experimental tasks may uncover misconceptions and maladaptive processing strategies that are inconsequential under normal circumstances, but could be detrimental in an emergency.

The ultimate goals of designing and training to enhance situation awareness depend on having satisfactory quantitative methods for measuring it. The development of measuring instruments, however, requires that the cognitive processes involved in achieving and maintaining situation awareness be understood. Several seminal papers have explored the information-processing demands on the commercial or military pilot (Beringer & Hancock, 1989; Endsley, 1988; 1989; Fracker, 1988; Regal, Rogers, & Boucek, 1988; Sarter & Woods, 1991; Whitaker & Klein, 1988).

The objective of this report is to expand upon this work by: (a) developing a theoretical framework based on psychological theory for understanding situation awareness, (b) using the framework to derive new strategies for assessing situation awareness experimentally, and (c) applying the proposed strategies to some sample assessment problems.

SITUATION AWARENESS AND PILOT ERROR

The problem of maintaining situation awareness in the modern, semi-automated cockpit has been brought to public attention through several dramatic inci-

dents involving pilot error. One such incident involved a China Airlines 747 jet near San Francisco:

The Taiwanese pilots, including a relief captain and flight engineer, had been flying nearly 10 hours, most of the time gazing at their instruments and watching the autopilot system navigate and fly the airplane. When the 747 went through a bit of turbulence, the auto throttle system, programmed to maintain a constant airspeed, first pulled the throttles back to reduce power from the engines and then, about 30 seconds later, increased the power to prior levels, just as it was supposed to. However, the outboard engine on the right wing did not respond fully, and despite the flight engineer's attempts to cope with the situation, the engine quit. The 747 has four engines, so the loss of one engine at high altitude is not even considered an emergency and is easily handled. Had the captain completely turned off the autopilot and followed correct procedures, the flight would have continued without incident....

The loss of the engine had caused the airplane to try to turn to the right. Sensing this, the autopilot began deflecting the control wheel in the cockpit to the left, attempting to stop the turn. With his attention focused, inappropriately and almost exclusively on the engine problem, the captain failed, according to the N.T.S.B. report, to realize that the airplane and the autopilot had become engaged in a vigorous tug-of-war. The airplane was trying harder and harder to turn, or roll, to the right because of the dead engine, and the autopilot was struggling to prevent this and keep the airplane level, as programmed.

Although the struggle between the two machines went on for nearly four minutes, the N.T.S.B.'s analysis afterward found, the captain was entirely oblivious to it because he was letting the autopilot fly and did not actually have his hands on the control wheel.... Finally, he disconnected the autopilot and took hold of the control wheel to fly the airplane himself. In that instant, the airplane immediately won the tug of war with the autopilot. The effect was similar to what happens in a real tug of war when one side, at a previously arranged signal, suddenly lets go of the rope... The 747 rolled dramatically to the right.... The roll continued and within a few seconds the 747 was on its back, plummeting earthward (Stockton, 1988, p. 63).

Fortunately, the incident ended without loss of life. The pilot recovered the plane at 10,000 feet, but not before the passengers had spent some time hanging upside down from their seatbelts. This incident provides an example of the problems that can occur when the transfer of control from automated equipment to human crewmember is carried out without appropriate awareness of the full situation.

HOW HAS SITUATION AWARENESS BEEN DEFINED?

The China Airlines incident described above helps to make the point that many factors are involved in achieving situation awareness. It is not surprising, therefore, that researchers have yet to agree on a definition of it. Consider, for example, the following two definitions:

[Situation awareness] means that the pilot has an integrated understanding of factors that will contribute to the safe flying of the aircraft under normal or non-normal conditions. The broader this knowledge is, the greater the degree of situational awareness (Regal, Rogers, & Boucek, 1988, p. 65).

[Situation awareness is] the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future (Endsley, 1988, p. 97).

These definitions, which are representative of those in the literature, differ in their emphasis on knowledge (i.e., product) and perception (i.e., process) respectively. Yet they both seem to capture the key features of situation awareness: an understanding of the meaning of events and the ability to anticipate the consequences of taking or failing to take particular actions.

A COGNITIVE VIEW OF SITUATION AWARENESS

We distinguish the process of achieving situation awareness, which is sometimes called situation assessment, from its product or consequences. The state of awareness with respect to

information and knowledge is the product. The process, in contrast, involves an active and dynamic series of cognitive activities. Maintenance of situation awareness is not easy because the process requires mental resources that may be in competition with ongoing task performance. The information gathering activities that contribute to situation awareness therefore may heighten workload momentarily. However, a principal benefit of achieving situation awareness is that the operator or crewmember is prepared to deal with upcoming events such that the extreme surges in workload that can occur in unexpected circumstances are avoided.

The development of suitable measurement techniques requires an understanding of the process by which situation awareness is achieved within the workload constraints imposed by the multiple tasks confronting the pilot. In broad perspective, although the crewmembers spend much of their time in routine, repetitive activities, a number of different, potentially knowledge-intensive and procedurally-complex tasks may demand attention at any moment. Each of these tasks is usually triggered by a stimulus event, such as a communication from Air Traffic Control (ATC) or a warning light, and, in order to obtain proper interpretation, may require that the crew initiate additional information-seeking behavior. The aircrew needs to be prepared for these tasks by having adequate understanding of the situation and knowing exactly what additional information is needed. For each such alerting signal or event, the crew must determine its relevance, its procedural and goal-related implications, and its urgency. Especially for the experienced crew, such events may often call forth highly practiced patterns and result in "automatic" responses that do not add to the workload. If not, however, assessing their significance may require access to the full range of human memory structures. The effort required to retrieve associated data and knowledge necessary

for deciding on a course of action necessarily results in an increase in mental workload. The framework we propose for understanding situation awareness draws on research and theory in cognitive psychology, particularly work on perception and cognition and on memory and attention.

NEISSEY'S THEORY OF THE PERCEPTUAL CYCLE

Neisser (1976) has formulated a view of the integration of perception and cognition that provides a useful way of thinking about situation awareness for several reasons. First, it resolves the tension between knowledge and perception that is evident in the above definitions of situation awareness. Building on Bartlett (1932) and Gibson (1979), Neisser argued that knowledge, in the form of schemas, or mental models, leads to anticipations of certain kinds of information and directs attention and exploratory movements to particular aspects of the available information (See Figure 1). The information that the perceiver picks up from the environment, or samples from the available information, in turn, modifies, or updates, what the perceiver knows about the immediate surroundings and influences what is known about the world in general. Situation awareness, in this framework, can be seen as both product and process. As product, situation awareness is the state of the schema. As process, situation awareness is the state of the perceptual cycle at any given moment. Situation awareness can be thought of as "the big picture" or context in which to interpret the flow of events. It allows the perceiver to attend to the right information at the right level of abstraction for the right task. It facilitates the handling and scheduling of multiple tasks and prevents overload. Strategies for reducing workload in multiple task situations will be touched upon in a later section.

Neisser's view of perception can also be expressed in terms of the new connectionist theories arising out of work on parallel distributed processing systems

(Norman, 1986). In these theories, newly perceived features are thought to be associated with other, pre-existing features through excitatory or inhibitory connections based on past experience with those features. Maintaining situation awareness, in these terms, is the optimal adjustment of weights, or tuning of the system, to anticipate the arriving data with minimal discrepancy.

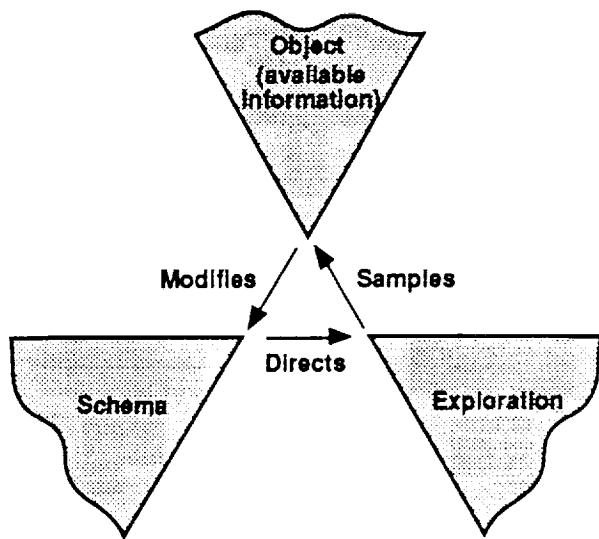


Figure 1. Neisser's (1976) view of the perceptual cycle. Copyright © by W.H. Freeman and Company. Reprinted with permission.

A second advantage to Neisser's approach to perception is that it emphasizes the role of meaning or comprehension that was mentioned in both of the above definitions. Following Gibson (1979), Neisser handles the problem of meaning by making it an integral part of the perceptual cycle.

The information picked up in vision is necessarily optical, consisting of patterns in the light over space and time. But optical information can specify objects and events at various levels of abstraction and meaning. When we perceive [a person's] mood, we are not engaged in the same perceptual cycle as when we are attending to his lip movements. We develop a different (though perhaps overlapping) set of anticipations; we pick up

information that extends over a different span of time (Neisser, 1976, p. 21).

A third advantage to this approach is that it handles the anticipatory processes that are mentioned in Endsley's (1988) definition and referred to as the ability to "think ahead of the aircraft" in Regal et al.'s (1988) discussion. In Neisser's perceptual cycle, the anticipation of events directs exploratory behavior. Exploratory behavior can occur in many different forms. Attention can be channeled to new objects with or without eye movements and with or without locomotion. Anticipation can occur in each of these situations. Anticipation can also occur at different levels in the continuum from perception to cognition. For example, at the basic perceptual level, anticipation means a readiness to perceive certain information, as when the optical flow specifies an impending collision or the occlusion of another aircraft by cloud cover. At the level of higher-order perceptual strategies, anticipation means focusing attention efficiently, as when an experienced driver (or taxiing pilot) focuses attention on a point at a distance from the vehicle rather than directly in front of it. At the cognitive level, anticipation means the consideration of possible outcomes. It includes knowledge-intensive activities such as contingency planning and diagnosis of system faults that may impact the flight. This level is best represented by Neisser's (1976) expanded view of the perceptual cycle (See Figure 2). The inner circle is the perceptual cycle shown in Figure 1. The outer circle is a more general exploratory cycle that includes actions taken to obtain information that is not present in the immediate environment. This kind of exploratory behavior is based upon knowledge of the world and its possibilities. At each of these levels, anticipation serves to help in the management of workload. Attentiveness to potentially important parameters during periods of low workload makes the perceptual process more efficient during periods of high workload, thereby protecting against overload.

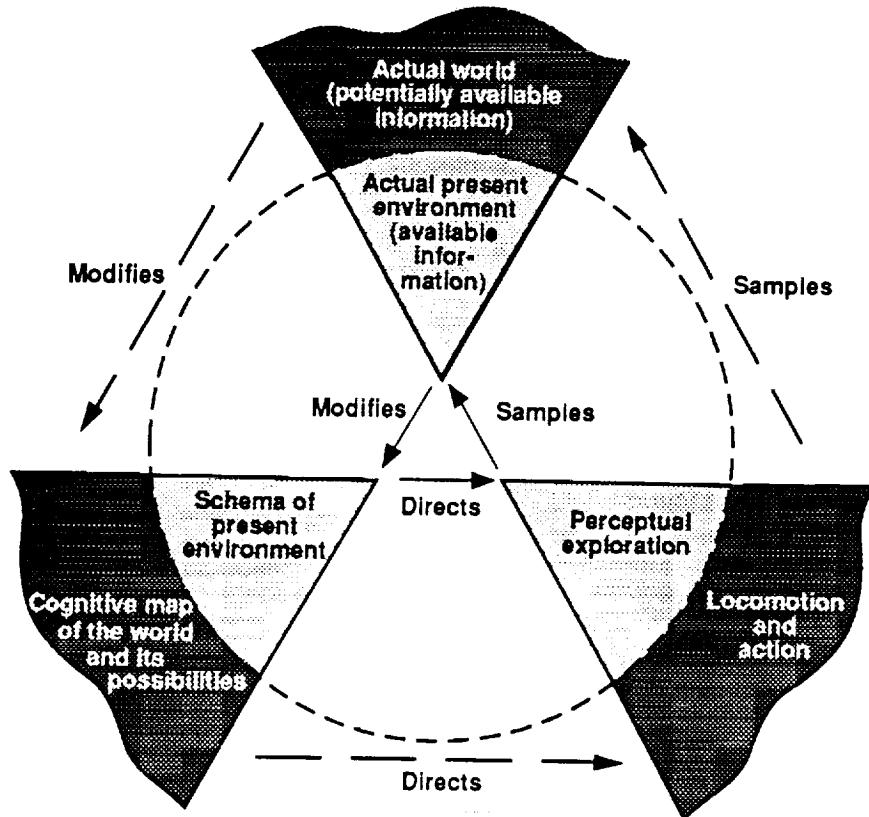


Figure 2. Neisser's (1976) expanded view of the perceptual cycle. Copyright © by W.H. Freeman and Company. Reprinted with permission.

REAL-WORLD CONSTRAINTS ON SITUATION AWARENESS

Having borrowed Neisser's notion of the perceptual cycle to clarify what we mean by situation awareness, we now turn to the important question of why it sometimes is and sometimes is not achieved. To answer this question, we need to consider the task that the pilot is trying to accomplish, the nature of the information that is available, the inherent constraints on human processing capabilities, and the nature of expertise.

The Pilot's Job

The most striking characteristic of the pilot's job is the multiplicity of tasks that must be accomplished. For example, at 2500 feet during an approach towards a rain storm, the pilot may have to perform the following:

- Monitor the descent (e.g., compare actual to desired flight path, airspeed etc.).

- Perform the prelanding checklist.
- Set the flaps/slats.
- Receive radio messages from ATC to reduce speed and watch for traffic.
- Enter the new altitude restriction into the Mode Control Panel (or equivalent).
- Look out the window for traffic.
- Respond to ATC about traffic seen or not ("seen" means pilot assumes responsibility for separation).
- Worry about the presence of windshear by looking for virga, lightning, or dust-rings.
- Watch airspeed, especially for evidence of a windshear encounter.
- Ignore radio traffic to other aircraft except where it contains warnings of nearby hazards.
- Monitor the copilot's performance.

Most of these tasks are triggered by specific events and many have a time frame during which they must be com-

pleted. Most require careful monitoring of information sources.

The Available Information

In carrying out these tasks, the pilot must be attentive to numerous sources of information. There is an ongoing stream of sensory information from the environment as well as cockpit visual and auditory displays, manuals, checklists, and communication from within or from outside the aircraft. Some of these information sources can be consulted at will, others intrude at unexpected times. For example, the following sensory events, all of potential significance, may occur in sequence while the pilot is in the midst of an approach:

- A flash of lightning.
- An altitude glitch that results from turbulence.
- Other signs of turbulence.
- A radio transmission.
- A speed reading that is right for flaps-down.

For each of these data or events, the pilot must determine relevance, procedural implications, and urgency. This determination can be more or less effortful, depending on factors that relate to human information-processing constraints. These constraints will be discussed in the next section.

Thus far we have described the multiplicity of tasks facing the pilot and the multiple sources of information to which the pilot must attend. Now we will consider the factors that affect the degree of effort that the pilot must exert to determine the significance of the incoming information. We will discuss two major factors: (a) the depth of the analysis required and, (b) the extent to which attention to the input requires a change in focus or context.

Depth of Analysis Required

In some circumstances, the pilot can choose to ignore or attend to incoming information on the basis of simple dimensions, such as location or modality. Location within the cockpit provides in-

formation about the function of certain inputs (e.g., primary flight displays are for control and guidance information, map is for navigation information). Cockpit location also provides information about the criticality of certain inputs (e.g., warnings are listed above cautions which are listed above advisories). Criticality is also specified by modality (e.g., the distinctive sound that signals a warning).

While the relevance of some inputs can be determined easily, the relevance of others is not signalled by either source or modality, but depends on its relationship to some particular aspect of the flight situation. For example, in an approach to landing, the significance of a radio message cannot be determined without further processing: a message from ATC to reduce speed and watch for traffic would certainly be relevant, a report from the preceding pilot would be relevant if it contained information about a potential hazard, a communication addressed to another pilot would be irrelevant, unless the pilot were nearby and it described conditions at the relevant airport, and finally, a radio weather report would be relevant if it pertained to the right part of the country.

While the significance of a warning or radio communication is relatively unambiguous, many of the inputs that pilots receive require additional information before their relevance can be determined. For example, a turbulence report en route to an airport could signal weather difficulties at the destination or merely a temporary disturbance. A simulation study involving precisely these circumstances showed that only the best pilots took the implications of the turbulence seriously enough to seek information about cross-winds at the designated airport; as a consequence, only the best pilots were prepared for the task of rerouting (Orasanu, 1990).

To summarize, one factor that influences whether and how easily an input can be processed has to do with the structure of the information. The significance of certain inputs is directly speci-

fied by superficial, attention-getting features while that of others requires deeper processing and further acquisition of information.

Change in Focus or Context

The depth of processing required to determine the significance of an input is an important factor in determining whether or not it will be properly attended. Recent work on memory and attention suggests that another factor affecting the processing of an input is the extent of the competing demands on the pilot's attention at any given moment. The bottom line is that human information-processing capabilities are not well suited to a multiplicity of simultaneous tasks. Thoughtful attention is modular. People can consciously think about only one thing at a time. As a result, they do not handle interruptions and distractions very well. Research has shown that even when an operator is faced with as few as two tasks and the "tasks" consist of nothing more than the detection or recognition of simple signals, the requirement to divide or switch attention between them may result in a significant loss in sensitivity or time that can be allocated to either (Broadbent, 1957; Schneider & Detweiler, 1988; Swets, 1984).

In addition, research indicates that mental shifts between topics or semantic domains require measurable time and effort and are prone to certain classes of biases and errors (Anderson & Pitchard, 1978; Bower, 1982; Sanford & Garrod, 1981; Schank et al., 1982). To the extent that incoming information is unrelated to the task in which the pilot is concurrently engaged, its interpretation must involve considerable mental workload. The more time and effort the pilot invests in its interpretation, the greater its potential for blocking notice or proper interpretation of other available data. The less time and effort invested in its interpreta-

tion, the greater the likelihood of misconstruing its implications. These ideas are captured in a theory about memory and attention that will now be described.

RECONCEIVING THE DEMANDS ON THE PILOT

A theory about memory and attention developed by Sanford & Garrod (1981) helps to clarify the cognitive demands on the pilot even though, interestingly, it was formulated to explain text comprehension. We will present the theory as it was originally formulated, in terms of the comprehension of events in a story, and then show how it applies to the comprehension of events in a flight.

Sanford and Garrod's Theory of Memory and Attention

Sanford and Garrod (1981) theorized that an individual's active memory consists of two bins, *explicit* and *implicit* focus. We find it convenient to think of these two bins as replacing the box labeled "schema of present environment" in Neisser's (1976) expanded view of the perceptual cycle (See Figure 3). *Explicit focus* corresponds roughly to what is conventionally labeled as "working memory." At any given moment, explicit focus contains a tightly-limited number of interrelated tokens of (or pointers to) larger knowledge structures in long-term memory. Although the contents of explicit focus are regulated more or less like a push-down stack, the maintenance of any given token depends not only on the recency with which it has been activated by the text or situation but also on its implicit relevance to the current interpretive stream.

Implicit focus, in contrast, subtends the full-blown representation of the situation that is partially represented in explicit focus. Information relevant to the knowledge in implicit focus cannot be brought to the interpreter's attention

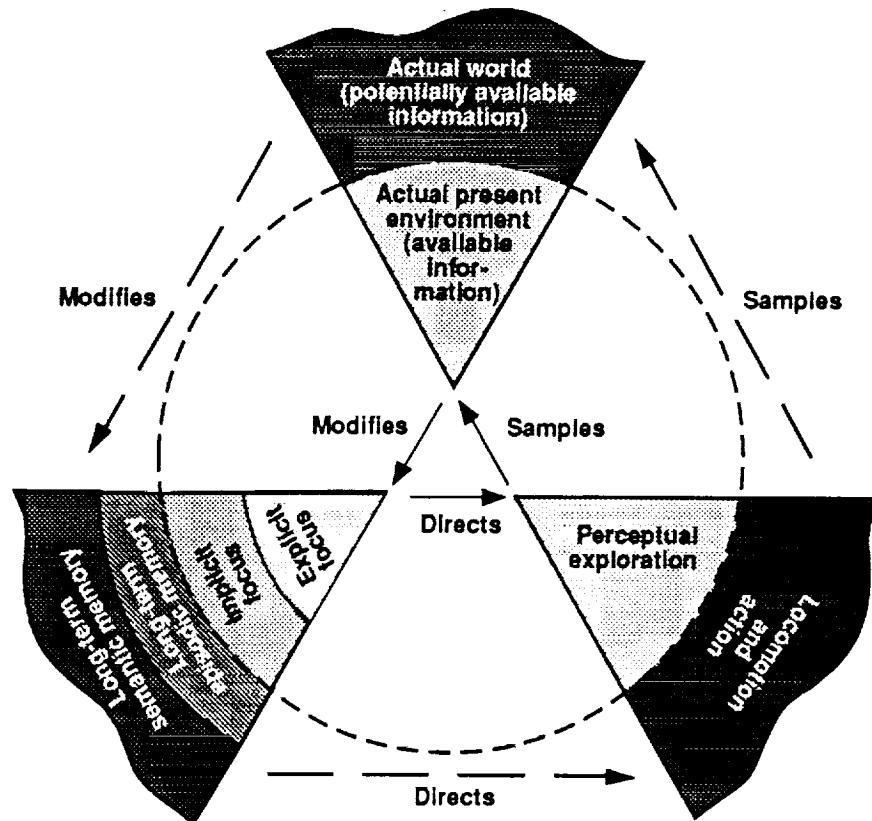


Figure 3. Modification of Neisser's (1976) view of the perceptual cycle

with either the speed or the obliqueness of reference that suffices for information in explicit focus. On the other hand, such information can be interpreted far more quickly and with smaller increases in workload than information that is unrelated to the contents of explicit focus.

To support these active memories, Sanford and Garrod suggest that the reader's *long-term* (currently inactive) *memory* is also sectioned into two bins. We think of these bins as replacing the box labelled "cognitive map of the world and its possibilities" in Neisser's expanded view of the perceptual cycle (See Figure 3). The first, *episodic memory*, contains a complete record of the knowledge structures that have been built or accessed in the course of reading the current text or, for our purposes, in the context of the current flight.

Meanwhile, *semantic memory* contains a person's lifetime accumulation of knowledge in general. Knowledge in either of these long-term memories can be brought to consciousness only given

considerable effort or strong cueing. Requirements to do so are thus expected to be relatively costly in terms of mental workload.

Extrapolating from the literature on text comprehension, we can anticipate some of the parameters that will control the ease or probability with which a given event is properly processed by the manager of real-world information. Because these ideas have not been tested or validated in the general domain of real-time information management, this extrapolation is speculative.

Events that are relevant to those aspects of a task on which a person is currently working should be readily assimilated because they will map themselves onto the knowledge currently in explicit focus; thus, for example, the pilot will readily notice and respond to changes in the glideslope indication that occur in the course of landing (See Figure 4). Events that pertain to the task but not to the particular aspect of the task with which a person is engaged are also expected to be

interpreted relatively quickly and cogently as they will map onto knowledge in implicit focus; thus, for example, even while tracking the glideslope, the pilot may be readily alerted to changes in engine noise that are consistent or inconsistent with landing experience (See Figure 4). In contrast, when the interpretation of an event requires consideration of knowledge in long-term memory, the probability or effort associated with its proper processing will depend on such factors as the transparency of its significance and the time available for working on it; thus, for example, very close to touchdown, the pilot will be relatively unprepared to receive and interpret unrelated communications.

Mission Categories

We find it useful to think of the tasks that command the pilot's attention as falling into the following mission categories:

- Macro-planning and -navigation.
- Local navigation, guidance and control.
- Communication outside the cockpit.
- Flight crew resource management.
- Cabin management.
- Routine management of physical equipment, resources and systems.
- Routine management of Flight Management Computer (FMC) and related crew-aiding systems.
- Troubleshooting of physical equipment, resources and systems.
- Troubleshooting of FMC and related crew-aiding systems.
- Bridging activities (breaks, teaching etc.).

Each of these categories has specific goals associated with it, although the goals in any one category may serve as subgoals for another category. The work by Sanford & Garrod (1981) on text

comprehension suggests further analogous constraints on pilot behavior. First, only one goal will be in explicit focus at a time. Data relevant to a new concern will more easily capture an operator's attention upon completion of a goal or subgoal. Although part of the goal structure may be available in implicit or explicit focus, the complete goal structure may be available only through long-term memory and thus require effort to retrieve. As a result, the necessity of switching to a task in another mission category before completion of the first task could lead to less than optimal deployment of attention on either task. Switching to another task within the same category will be less disruptive because the two tasks are likely to have a common goal and common implicit information.

The effort required to switch attention between tasks in different mission categories means that the original task may be forgotten altogether, since the goal is no longer in working memory, especially if further distractions prevent the pilot from focusing on the goal structure long enough to reactivate the original task from long-term memory. Such memory lapses can be dangerous.

An Airline Accident

Many of the airline accidents that have been reported lend credence to the above depiction of the limits on human attention. An example is provided by a recent accident at the Detroit Metropolitan Airport (NTSB, 1988). The aircraft took off without setting its flaps and crashed. Although the crew had begun the pre-flight checklist properly (mission category: management of physical equipment, resources, and systems), they were interrupted by Air Traffic Control before verifying the

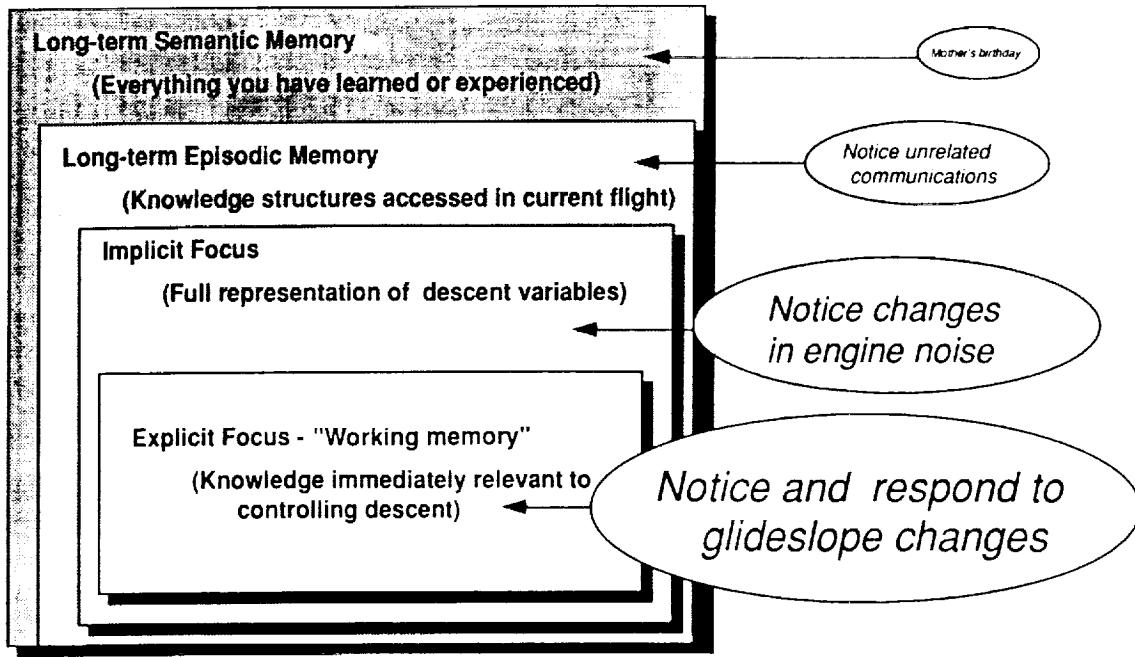


Figure 4. Hypothetical view of the pilot during landing. The boxes show the contents of memory; the circles show the objects of awareness.

status of the flaps (mission category: communication outside the cockpit). Although they might still have resumed the checklist routine prior to take-off (management of physical equipment), other issues usurped their attention: they were confused as to which taxi-way to use (local navigation), the runway direction had just been changed, and weather and runway conditions were not provided until the aircraft was already taxiing (macro-planning).

Although, with proper handling, the aircraft could have become airborne without flaps, the crew had been given a windshear alert. When the problem with the flaps expressed itself during take-off, the symptoms were interpreted -- and responded to -- as though they were caused by windshear (See Figure 5). As Figure 5 illustrates, this misinterpretation is understandable in light of the salience of windshear in the pilot's memory and the similarity of symptoms in the two cases. Very likely, the cost of this misinterpre-

tation was compounded by what Norman, (1986) calls a tendency toward "cognitive hysteresis," the fact that human information-processors are notoriously bad at seeking disconfirming evidence once a hypothesis has been partially confirmed. Evidently the distractions from the competing tasks left few resources with which to counteract this natural tendency to stand by an initial decision. The Detroit incident, in short, is typical of accidents resulting from human error. A series of minor errors, none of which was fatal in itself, together resulted in a disaster (Reason, 1990).

THE NATURE OF EXPERTISE

Next we briefly consider how experienced pilots are able to perform the multiplicity of tasks without error, in spite of inherent limitations on the human information processing system. The limitations are not so debilitating as they might

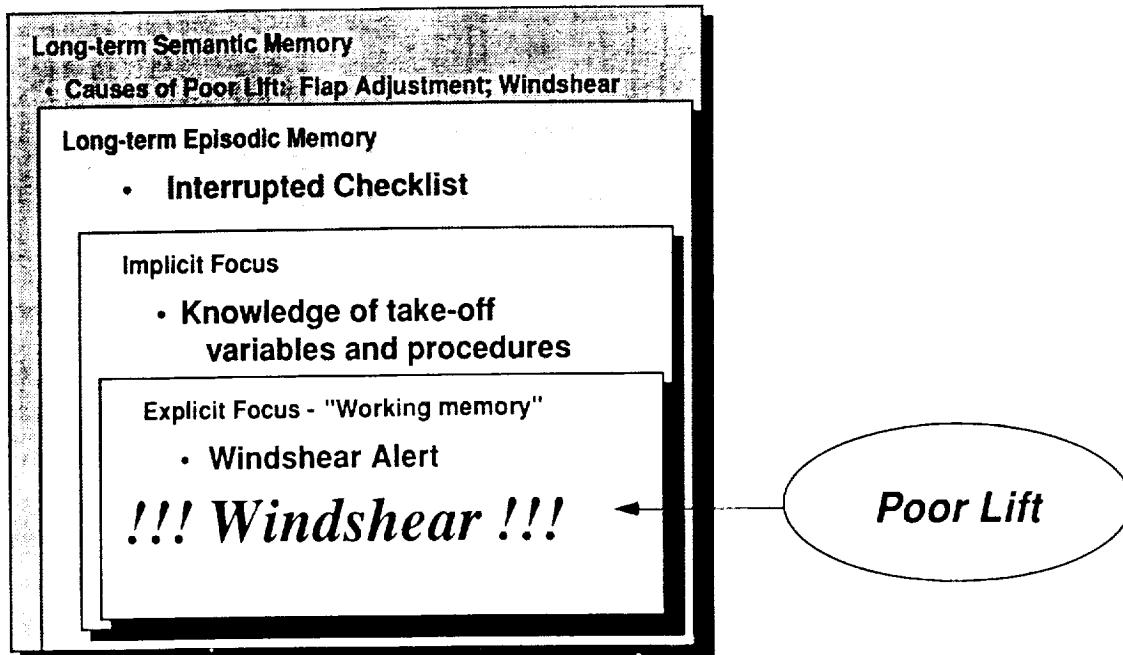


Figure 5. Hypothetical view of the pilot during the incident described in the text. Expectations based on a windshear alert caused the pilot to overlook the true cause of the lift problem.

first appear because of the extreme adaptability and resourcefulness of the human cognitive system. Schneider & Detweiler (1988) have identified seven types of compensatory strategies that people use to overcome performance limitations in multi-task situations. These strategies are thought to emerge routinely in the course of practice on multiple tasks, although the first two may be specific to conditions of high workload (Schneider & Detweiler, 1988).

Shedding, Delaying, and Preloading Tasks

The first set of strategies consists of the following:

Experience in performing multiple tasks enables one to anticipate and monitor the consequences of delaying or eliminating a task (Schneider & Detweiler, 1988, p. 556).

In the China Airlines accident, the pilot was preoccupied with looking out the window at the defective engine with the consequence that he delayed disconnecting the auto pilot, with disastrous results.

Preloading involves preprocessing information prior to the onset of the critical workload segment (Schneider & Detweiler, 1988, p. 556).

One aspect of preloading, contingency planning, has been found to be characteristic of expert flight crews. Orasanu (1990) found that air crews who were rated as most successful in a simulated flight study (i.e., made the smallest number of errors) differed from crews rated less successful in the extent and timing of their contingency planning.

Successful crews were likely to make plans (e.g., regarding weather closing of airports) and did so well in advance of when they were needed. The less successful crews, by contrast, were not prepared for the rerouting and ended up having to make plans during a peak workload (e.g., while having to deal with a simultaneous hydraulics system failure).

The picture we get here is of good Captains who are planful, anticipate difficulties, use time during the normal phase to prepare for higher workload periods, and who are ahead of the curve. They stop and

think, then get information they need and carry out their plans. Emergency actions are rehearsed so the crew is primed to go into action when the time comes (Orasanu, 1990, p. 13).

Letting Go of Unnecessary High-workload Strategies

The expert finds short-cuts that reduce the workload. This aspect of pilot performance is particularly important when problems arise. The less effort required to monitor the flight, the more likely it is that the critical monitoring function will be maintained when the pilot has to deal simultaneously with a system failure.

Utilizing Non-competing Resources

This strategy involves learning to allocate task components to minimize resource competition. The experienced pilot, for example, would be wary of carrying out another task that used spatial strategies while navigating, but would feel free to engage in activities that used quantitative strategies, such as calculating flying time. Successful cockpit resource management strategies also fall into this category.

Multiplexing Transmissions Over Time

This strategy has to do with optimizing the division of attention across tasks:

[Researchers] have shown that after extended training, human operators learn to sample instrument gauges at the optimal rate based on the relative information rate of each channel. The allocation of internal control processing may be tuned through experience in a manner comparable to the way the operators allocate attention among gauges (Schneider & Detweiler, 1988, p. 558).

Shortening Transmissions

This strategy consists of finding the minimal time that can be spent in looking at a gauge or source of information without sacrificing accuracy. Shorter viewing times also result, with experience, from learning to ignore

irrelevant features and from chunking (See below).

Converting Interference from Concurrent Transmissions

This strategy involves finding ways to reduce interference from competing sources of information, by deliberate shifts of attention. Tuning out words spoken by your fellow crewmember when ATC is contacting you about something else is an example. The crewmember's message may be retransmitted at a later time.

Chunking Transmissions

This strategy consists of grouping information into related clusters to overcome the limits of working memory. Work on expertise in many domains has shown that the expert is more likely than the novice to perceive the higher order relations or patterns in the information that permit chunking. In a replication of an earlier work on chess masters, Chase & Simon (1973) found that experts, who were asked to reproduce the position of chess pieces from a real game after a brief glance, remembered more board locations than did novices. Their performance was not better than that of novices, however, when the pieces were randomly placed on the board. Thus, the experts were able to encode legitimate board arrangements into higher order units that could be perceived in a glance. In terms of flight management, we would expect more experienced pilots to differ from those with less experience in the ease and accuracy with which the patterns of information (i.e., "situations") derived from the displays could be perceived and remembered.

SUMMARY

By way of review, Table 1 presents what we feel are the key elements of situation awareness discussed thus far. The table is organized into three possible classes of situations faced by the pilot: routine, non-routine, and emergency. These classes all involve anticipating or "thinking ahead of the aircraft." This

ability becomes increasingly important as the situation changes from a routine flight to an emergency. Anticipation is accomplished in routine situations through the perceptual cycle depicted in Figure 1. Emergency situations, which require contingency planning and diagnosis, are accomplished through the larger exploration cycle, depicted in Figure 3. For each of the situations in Table 1, we specify the components that can require attention and the benefits of awareness. We then break down each of the situations into the relevant mission categories and summarize differences between novice and expert performance.

STRATEGIES FOR THE MEASUREMENT OF SITUATION AWARENESS

In this section we address the problem of assessing situation awareness. This problem has been discussed in several important papers (Beringer & Hancock, 1989; Endsley, 1988; 1989; Fracker, 1988; Regal, Rogers, & Boucek, 1988; Sarter & Woods, 1991; Whitaker & Klein, 1988). The techniques that have been emphasized thus far are the probe technique and think-aloud protocols. In the probe technique, a simulation is expectedly or unexpectedly interrupted and the crewmember is asked to report on the state of various classes of information reflective of the state of his or her aircraft (Endsley, 1988; 1989). In think-aloud protocols, the crewmember is asked to provide a running narrative explanation of what she or he is doing during the execution of a scenario or during a post-flight debriefing. The need for reliable, sensitive measures suggests that we exploit these conventional approaches, but that we also go beyond them in order to broaden the scope of measures and to capture more of the properties of both the process and product of situation awareness.

In the remainder of this section, we discuss techniques for assessing situation awareness. The idea is to provide a range of strategies that can be fitted to different needs. The application that concerns us here is the evaluation of the

situation awareness afforded by proposed new cockpit aids. However, we believe the strategies that we are proposing are general enough that they can be applied in areas other than aviation and to other assessment needs, such as the evaluation of the effects of training programs or the development of certification requirements.

The strategies we will describe for measuring situation awareness fall into three categories:

- *Measures derived from scenario manipulations:* Situation awareness is inferred from performance on specially designed scenarios that place a demand on crewmembers' attentional resources and are hypothesized to require a sufficient state of awareness for good performance.
- *Direct measures:* Situation awareness is inferred from a variety of probes and experimental tasks that tap the crew's knowledge and strategies during a scenario or following a brief vignette
- *Model-based measures:* Situation awareness is inferred from the performance of a computer program that models the actions, knowledge states, and information processing strategies of a human (i.e., a flight crewmember) for a given situation.

Although, we will discuss each of these measurement categories separately, they are intended to be used in combination.

MEASURES OF SITUATION AWARENESS DERIVED FROM SCENARIO MANIPULATIONS

We believe that with proper scenario design, much can be revealed about the situation awareness of the crewmember through performance measures alone. The key to designing scenarios for this purpose is to manipulate the demands on attention. Under normal circumstances, performance may not be a good indicator of how easily the pilot can stay abreast of the situation. A scenario in which

Table 1
Characteristics of Pilot Situation Awareness

| Situation | Situation Components | Benefits of Awareness | Mission Categories | External Information Sources | Novice | Expert |
|--------------------|-----------------------------------|--|--|--|--|---|
| Routine | Spatial orientation | Mid-air collision avoidance | Local navigation, guidance and control | Sensory information from the environment | Expendng unnecessary effort | Utilizing non-competing resources |
| | Environment | Terrain avoidance | Communication outside the cockpit | Cockpit visual and auditory displays | Not perceiving patterns | Multiplexing |
| | Routine goals | Robust decision making in the face of: | Flight crew resource management | Extra- and intra-aircraft communication | | Shortening transmissions |
| | Procedures for attaining goals | | | | | Converting Interference |
| | Aircraft system status | Turbulence Cross winds Windshear Loss of visibility | Cabin management | Recorded flight plans | | Chunking |
| | Aircraft performance | | Routine management of physical equipment, resources, and systems | Flight management computer | | |
| | Crew responsibilities & knowledge | | Routine management of FMC and related crew aiding systems | Flight manuals and checklists | | |
| | | | Bridging activities | Future 3D navigation aids | | |
| Non-Routine | Special constraints | Improved decision making | Macro-planning & navigation | Future route diversion aids | Making last minute plans during high workload | Shedding, delaying, and pre-loading tasks |
| | Contingency plans | In the face of: Go around Weather rerouting | | | | |
| | | | | | | |
| Emergency | Unusual symptoms | Improved fault management | Diagnosis of physical equipment, resources, and systems | Future fault finding aids | Fixating on one or two salient possibilities | Letting go of high workload strategies |
| | Trouble-shooting techniques | | | | | |
| | Emergency procedures | | Diagnosis of FMC & related crew-aiding systems | | Ignoring vital flight information while trouble-shooting | |

Note: The rows are cumulative. In other words, the entries for emergency situations include all the entries for routine and non-routine situations.

events are entirely predictable and the effort required to assess the situation is minimal is not likely to be sensitive to differences in situation awareness. Therefore, scenario variations are needed that place competing demands on attention. Attentional demands can be increased in several ways (See section on real-world constraints on situation awareness). Two that are relevant for the design of scenarios are:

- Increase the depth of processing required to comprehend the inputs.
- Arrange for a shift in attention or a change of focus.

These factors give rise, respectively, to the following strategies for assessing situation awareness.

Design Scenarios with Subtle or Misleading Indicators

The depth of processing required to interpret an input in a scenario can be manipulated by introducing either conflicting, potentially misleading, or very subtle indicators. Situation awareness, or the lack thereof, can be inferred from the nature and timing of the choices made. The idea is that instruments and aids that provide good situation awareness will enable the crewmember to focus on subtle, but critical information and avoid being led down the "garden path" by information that is potentially relevant, but turns out to be inconsistent with the overall pattern. This approach was successfully applied to the evaluation of displays in a nuclear power plant control room by Woods, Wise, and Hanes (1982).

An example of a scenario with subtle indicators is provided by Orasanu (1990). In one of the scenarios in her study, information about crosswinds at the airport provided definitive anticipatory information about a weather closing at the destination airport. This information was important because it allowed the crew to plan ahead while the workload was low. It was subtle, however, because it had to be solicited explicitly on the basis of another, less

definitive indicator, turbulence en route. Only half the flight crews in her study solicited this information. In this example, the decision to seek information about crosswinds would be a good index of situation awareness.

The Detroit accident, discussed earlier, is a clear case in which ambiguous information, concerning windshear, led to difficulties. Another is provided by the Midlands plane crash (NTSB, 1989). In this tragic incident, the pilot misinterpreted the indicators concerning an engine problem and turned off the good right engine instead of the faulty left engine. There were at least two indicators that helped to mislead the pilot, although there were other indicators that would have clarified the situation. First, there was smoke in the cockpit and smoke is usually associated with a problem in the right engine because of venting patterns. Secondly, there was a vibration that coincidentally ceased after the right engine was turned off, suggesting that the malfunction had caused the vibration. One could construct analogously ambiguous experimental scenarios and use the time and nature of subjects' decisions, e.g., to turn off an engine, as an index to their situation awareness.

Design Scenarios with Interruptions

Demands on attention can be manipulated in a scenario by interrupting the primary task for a period of time. The degree of disruption can be observed through the quality of the decisions that are made as a result of having to divide attention and the time needed to restore the original context after the disruption. The notion here is that cockpit aids that increase situation awareness should help the crewmember to deal with the competing demands on attention occasioned by an interruption.

The Detroit accident again provides a case in point (See Figure 5). Interruptions from ATC while the crew was working on a checklist resulted in a failure to return to that task and, consequently, a failure to set the flaps -- and

apparently to recall that the flaps had not been lowered -- before take-off. This oversight had severe consequences because other factors conspired to make the crew believe the problem was windshear and prevented them from taking the correct actions.

Although it is not difficult to introduce realistic interruptions into a scenario, for purposes of assessment it may be sufficient to interrupt the ongoing flow of information artificially by blanking the cockpit displays. With this paradigm, both the actions taken while "blind" and the effort required to restore the original context when the display returns can serve as an indicator of the extent to which cockpit systems have supported the development and updating of the crewmember's mental model or "cognitive map." Further, by altering certain data on "restoration," we can make finer assessments of the scope and focus of subjects' awareness.

DIRECT MEASURES OF SITUATION AWARENESS

A number of more direct measures of situation awareness can be obtained by supplementing the scenario observations, discussed above, with experimental tasks and probes that shed light on the cognitive processes underlying the crewmember's performance. We will describe some assessment techniques that have proven to be robust within the field of experimental psychology and could be adapted to the problem of assessing situation awareness. The four techniques that will be discussed are:

- Queries.
- Recall of inputs.
- Category judgments (i.e., normal/abnormal).
- Information seeking/eye movements.

Queries

The technique of periodically querying the crewmember about situational information, either during the scenario or at the end, is the most commonly used method for assessing situation awareness (Endsley, 1988; 1989). While this is a

useful technique with high face validity, there has been little discussion about how to select probes for different situations. We would advocate a systematic approach to the use of probes that includes questions about information and indicators that vary systematically in their relevance to the task in immediate focus and to the overall status of the aircraft and flight plan. These would include spatial orientation, environment (weather, traffic), current goals and procedures, contingency plans, aircraft systems status, aircraft performance, diagnoses of malfunctions, and crew responsibilities and knowledge (See Table 1).

Recall of Displays

Another strategy for measuring situation awareness is to test the ease with which the relevant details of a display can be committed to memory. The rationale for this technique comes from studies of expertise in a number of domains, most notably chess (Chase & Simon, 1973). The findings from these studies consistently show that experts have a better memory for the details of a display (e.g., the positions of playing pieces on a chess board), after a brief viewing, than do novices. The consensus from this research is that experts are able to perceive meaningful patterns in the display that reduce memory load. This conclusion has been supported by the additional finding that experts are no better than novices at remembering incoherent displays (e.g., randomly placed chess pieces).

Because recall is strongly affected by meaningfulness, and the perception of meaningful events is the essence of situation awareness (Vicente, 1988), intentional recall of relevant displays can serve as an index of situation awareness.

Category Judgments

This technique involves measurement of the time taken to judge display inputs as "normal" or "abnormal." The rationale for this measure is similar to that for the recall measure. The same

skilled perception (i.e., situation awareness) that allows for "chunking" of the information, also should enable the perceiver to notice deviations from the expected pattern. We envision this technique being used with short vignettes that terminate with a frozen display. The viewer's task would be to rate the situation as normal or anomalous as quickly as possible. Normal displays would contain readings that would be likely to occur. An abnormal display could be constructed for each normal display by changing one or more of the indicators to a value that could not occur in a real situation.

Information Seeking

The last strategy consists of techniques for determining the information to which the crewmember is attending at any given time. For example, it can be used to determine whether and why a crewmember failed to look at the critical portion of a display in a scenario. One way of accomplishing this goal is to restrict the available information sources to those that the person actively seeks. In other words, no information from displays, communication sources, or flight aids would be presented unless it were specifically requested. The crewmember would be instructed how to obtain the needed information, perhaps through the use of a menu or touch screen. A possible disadvantage of this technique is that the viewer might fail to solicit information that would have been attended had it been present. A more sophisticated, but labor-intensive way of accomplishing the same goal would be to record eye-movements during the scenario.

MODEL-BASED MEASURES OF SITUATION AWARENESS

We are proposing the use of model-based measures as a supplement to the scenario-based and direct measures discussed above. The scenario-based measures allow us to create circumstances in which situation awareness is particularly important and observe whether the crew's response is reflective of that awareness. The more direct measures

allow us to observe those aspects of situation awareness on which the crewmember can report or that can be inferred from specifically designed performance measures. We propose the use of model-referenced performance measures in an effort to probe deeper into the state of the cognitive structures that support effective situation awareness.

The kind of human-machine performance model to which we are referring includes: (1) a scenario driver to set the context in which the behavior is to be measured, (2) a representation of those aspects of a system being controlled that impact on the human performance requirements, (3) a representation of the goals and tasks that the human machine-system team is required to perform, including descriptions of the procedures by which they are performed and (4) representations of primitive human performance capacities and limitations, such as visual scanning rules, the state of information in memory, methods for assigning task priorities, and resource-based workload limits.

Examples of such models are the PROCRU model that describes the behavior of the pilot-flying and pilot-non-flying of a Boeing 727 accomplishing a final approach and landing (Baron, Zacharias, Muralidharan, and Landcraft, 1980), and the AIRT model that describes the behavior of the controllers in an Air Force forward air control center. (Corker, Cramer and Henry, 1990). While human performance models of this kind have been and are being developed, the approach we are suggesting is speculative, since they have not previously been used in support of real-time performance analysis in tandem with an on-going simulation test of human performance. However, we believe the state-of-the-art has advanced to the point that exploring such an application is now a productive research strategy.

To apply models for the assessment of situation awareness in flight deck simulation studies, one would build a model of the flight deck control and display environment, of the behavior of the

aircrew that was valid for the flight crew tasks under study, and a simplified model of the aircraft performance. The model would be run in parallel with the actual aircrew participating in the simulation trial. It would be set up with the same scenario background and flight plans. Information received by the human crew would also be input to the modelled crew. This information would include data about the state of the aircraft and its systems that is available from flight deck displays and the content of ATC communications. Modelled crew actions would affect the modelled aircraft state.

The modelled system would be regularly up-dated with information from the actual on-going flight simulator trial to insure that the modelled system did not get out of synchronization with the real-time performance of the human crew in the simulated aircraft. To the extent that the modelled system is a valid representation of the real system, then the crew behavior will mimic the behavior of the human crew.

When a successful mimic is achieved, then the analyst has available in the model a representation of the internal state of the crewmember at each moment in time. This representation is exactly what is required to undertake an in-depth analysis of the state of the crewmember's situation awareness. At critical instants one could examine the current contents of the crewmember's memory representation; one could inspect and evaluate the priority ordering of the queue of tasks waiting to be accomplished. Tracking the contents of memory and task execution protocols would provide clues to points where lack of situation awareness is due to excessive monitoring requirements or task loading. Similarly it will highlight the source of differences in awareness that can be attributed to different display configurations or approaches to automation. In short it can provide substantial augmentation to situation awareness data without disrupting the conduct of the live

simulation itself or relying on the self reports of crew-member subjects.

ASSESSING THE IMPACT OF NEW TECHNOLOGIES ON SITUATION AWARENESS: TWO EXAMPLES

To conclude, we present two examples of how these strategies for assessing situation awareness can be used to evaluate the effect of proposed cockpit aids. The first example concerns a panoramic display that is intended to increase spatial awareness during landings. The second concerns a fault finder designed to increase the aircrew's awareness of system malfunctions.

EXAMPLE 1: EVALUATION OF PANORAMIC FLIGHT PATH DISPLAYS

An evaluation of a proposed new cockpit display that provides a panoramic, 70 degree field of view of the landing strip under instrument flight conditions is currently being undertaken by Russell Parish's group at NASA-LaRC. The display depicts the intended flight path as a row of goal posts along an imaginary pathway in the sky. A perspective view of a series of rectangles shows the position of the plane with respect to the intended flight path. At a point when the landing strip would come into view out the window, it is represented on the display in its proper perspective. Other aircraft operating in the near-by airspace are also indicated and are enclosed in an attention-getting red-outline box if they become a collision threat.

This panoramic display is being compared to the use of conventional instruments, including a plan-view navigation display and conventional attitude display. Special scenarios consisting of a final approach and landing are being used for this purpose. The investigators are interested in how each of the displays affects the pilot's ability to maintain spatial awareness and control of the aircraft in a range of circumstances. We describe their assessment design here because it is illustrative of many of the strategies we have suggested.

Introduce Subtle or Misleading Information

In some scenarios, a difficult flight path with several turns will be required. At the most difficult point in the flight path, another aircraft in the vicinity will unexpectedly turn onto a collision course with the aircraft under control of the pilot under study. The measure of situation awareness will be the time to detect the offending aircraft and the nature of the actions taken to correct it.

Introduce Interruptions

During other scenarios, the screen will go blank for five seconds and the plane will return to an unexpected location somewhat off the flight path. The pilot will be required to assess the new location and execute suitable maneuvers to return to the intended flight path. The measure of situation awareness is the time needed to initiate corrective action and the nature of any maneuver errors committed in returning to the intended path. It seems likely that the panoramic display will greatly facilitate this task, but quantitative measures of the differences in display conditions will be illuminating. Another potential measure would be the degree to which subjects' maneuvering of the plane during the time the screen was blanked conformed to the original flight path. Close conformity under the panoramic display condition would suggest that the display had fostered the development of a good mental model that allowed for accurate prediction of the intended flight path.

Queries

Direct measure of spatial awareness will be obtained by interrupting some of the scenarios at unexpected points and asking subjects questions about the location of their aircraft and other aircraft in the area. The evaluation of spatial awareness, as illustrated by this study, is perhaps the most straightforward application of our approach. Yet it has required considerable creativity on the part of the evaluation team. The assessments

of situation awareness outside of the spatial domain, we suspect, will require even more ingenuity. For purposes of discussion, in the next example we consider how we would go about evaluating situation awareness regarding system malfunctions.

EXAMPLE 2: FAULT-FINDER

Much thought is going into the design of fault-finding aids. The Midlands Airline accident (NTSB, 1989), in which a crewmember erroneously turned off the good, right engine instead of the defective, left one causing the airplane to crash on landing when the left engine failed, has provided a cogent example of a case where such an aid might have been advantageous. Although the aircrew consulted the relevant instruments before making the decision, they apparently did not interpret them correctly. For purposes of this discussion, we consider how we would go about comparing some of the new aids that are being developed to enhance awareness of system malfunctions (e.g., Abbott, 1990; Hudlicka, Corker, Schudy, & Baron, 1989). These aids perform one or more of the following functions: indicate whether instrument readings for various system parameters are within the expected range, produce hypotheses about possible malfunctions, and suggest procedures for dealing with the malfunction.

We would first design some scenarios in which a malfunction occurred and observe crew performance with the different flight aids. To increase the sensitivity of our measures, we would introduce factors into the scenario that increase the difficulty of attaining situation awareness:

Introduce Subtle or Misleading Information

The Midlands accident (NTSB, 1989) provides some useful suggestions. The accident report points to problems of misleading and subtle information. Misleading information (the coincidental cessation of vibrations and smoke when

the good engine was turned off) caused the crew to think they had turned off the defective engine and to disregard the instrument readings. Subtle information (the 16 degree roll to the left after the autopilot was turned off but before the right engine was throttled back) was not noticed by the pilot, although it should have provided an indication that the left engine was not functioning properly.

Introduce Interruptions

We would again turn to the Midlands incident for inspiration. The flight crew in that scenario was diverted by the following string of interruptions: the need to communicate with the passengers via the cabin address system, a weather information advisory, the need to reprogram the flight management system, ATC messages concerning a new radar heading, further descent clearance and a new radio frequency, checklist requirements, a radio call requesting that a test call be placed to the aerodrome fire service, and distracting radio calls from other aircraft on the same frequency. These interruptions may have played a significant role in the incident because they prevented the pilot from returning to a conversation with the copilot in which he was just beginning to question the evidence that had led them to shut down the wrong engine.

Second, we would use several kinds of direct measures of situation awareness, as discussed above:

Queries

We would ask questions to probe the crewmember's understanding of various aspects of system functioning either during or after the scenario. The accident report from the Midlands incident, in its thoroughness, provides a model of how to accomplish this task. In the aftermath of that incident, information was obtained from the crewmembers about: the meaning of the indicators -- how reliable the readings are on this and other aircraft (this question was important because vibration readings which were accurate in this aircraft had been unreliable on other

aircraft); the meaning attributed to various symptoms, such as smoke in the cabin (this question was important because this aircraft was different from many others, in which only the right engine ventilates into the cabin); or the failure of the right engine to restart when it was engaged at the last minute (the engine start levers, or fuel valves, were found in the cutoff position after the accident).

Recall of Displays

One problem with the instruments for indicating engine malfunctions that became clear in the Midlands accident is that it was difficult to perceive patterns. Recall of the information in briefly presented displays, therefore, would be one measure that would be important for assessing the effectiveness of the different aids.

Category Judgments

Another way of assessing the meaningfulness of the various displays would be to present displays of information for both right and left engines, where either one, both, or neither was defective, and have pilots make judgments as to which were abnormal, as rapidly as possible.

Information Seeking or Recording of Eye Movements

Finally, we would supplement the above measures, if possible, with the recording of eye movements to ascertain how frequently and how early the aids were consulted and how rapidly information was extracted from them.

AUTOMATION AND SITUATION AWARENESS

As has been suggested, one of the motivations for measuring situation awareness today is the interest in making the introduction of new cockpit aids and automation concepts more human centered. In this section we elaborate on the problems that automation can introduce and the potentially new situation awareness requirements it invokes.

In the past the pilot obtained situation awareness by doing, by being directly involved in every aspect of the flight. Information that supported situation awareness was a combination of that which the pilot was compelled to seek in order to be able to perform a task adequately and the feedback that occurred as a result of the pilot's actions. Situation awareness was enhanced by the lack of complexity. The danger, however, was that the manual workload of controlling the aircraft could be so high under certain conditions that the extra attention that was required to construct good situation awareness was not available.

Automation can, and has relieved the flight crew of many functions, thus lowering selected aspects of workload. It could be argued that by using automation the pilot has more time to actively seek the information that allows him to be situationally aware. As we have argued in previous sections, achieving situation awareness is an active process of seeking, prioritizing and interpreting information. Successful accomplishment incurs a workload penalty; someone who has no spare workload capacity because he is manually controlling attitude, speed, etc., will not have the same ability to achieve situational awareness as someone who has off-loaded part of the manual workload to automation.

However, this eliminates the apparently natural ability to construct situation awareness "by doing". As automated systems have been introduced, the pilot has become more and more removed from certain aspects of the flight. He no longer is privy to all the raw data required to adequately perform a task when that task is performed wholly or in part by the automation (even when he is responsible in the case of a failure). By the same token, the feedback given to the automation, which it requires in order to assess its dynamic performance, and the inputs the automation must make to compensate for a changing situation, are not always passed along to the pilot because they are considered irrelevant. However, they can be highly relevant

when the pilot is asked to take over on short notice in the event of a failure.

In addition, the introduction of automation adds many more information elements and interactions that one must keep track of. That is, each new piece of automation becomes part of the situation of which one must be aware. When a pilot manually controls the attitude of the aircraft, he must be aware of the chain of effects of stick movement, control surface movement, and the dynamics of resulting attitude changes. When the autopilot is used, the pilot should still be aware of these items and often is not. Further, in the case of an unexpected event, he should be aware of the characteristics of the automation itself, such as: current autopilot mode, if there is more than one; factors or actions that will cause it to disengage; the values that drive it; and what happens if it fails. This all puts additional situation awareness requirements on the flight crew, which technology-driven automation often does not make salient and when it does, it is at times when the human pilot is least able to cope with workload demands of the unexpected event.

As a result, any kind of automation that requires direct crew interaction in either routine or back-up mode operation expands the situation of which the pilot must be aware. It adds more items to the queue of information that must be tracked and it adds layers of complexity or abstraction to the situation so that the workings and nature of the remote layers are less observable.

Automation that is to be human centered must use the information manipulation, communication and presentation capabilities associated with it to provide informational support and feedback about the state of the automation itself in an expeditious manner. In the ideal case, this cannot only compensate for the loss of direct feedback and simplicity, but can provide the pilot with perspectives that were unavailable without the automation (See Vicente & Rasmussen, 1990 for an example in the domain of process control). Human-centered automation offers

the potential of greater pilot situation awareness than either no automation or technology-driven automation, which, in turn, should lead to improved overall human-system performance and aircraft safety.

CREW AWARENESS

The focus of the discussion of this report has been on the situation awareness of an individual crewmember, the pilot, the co-pilot or where applicable the flight engineer. There is also beginning to be a literature on the topic of crew awareness by which we mean the awareness of the aggregate set of individuals responsible for the safety and effectiveness of the flight (Wellens, 1989; 1990). Flight procedures identify that some features of awareness are required of the pilot-flying while others are required of all crewmembers.

Assessing crew awareness focuses attention on the communication processes among crewmembers. Information on flight instruments is available to both crewmembers; however at critical moments one is usually attending to out-of-cockpit sources while the other is "flying the instruments". When one is completing a check list, it is important that the other be aware of non-routine findings when they have a potential impact on current or future decisions. A lack of crew awareness may have played a role in the Midlands incident as well. Evidently, the flight attendants who saw flames coming out of the left engine were unaware that the pilot thought the right engine was the defective one. To make matters worse, the pilot was unaware that the flight attendants who had seen the flames in the left engine, had not heard his announcement. As was discussed earlier, experienced and effective crews discuss precautions, although they may never need them, while workload is low so that they have little need for discussion when a problem arises. Less effective crews, by contrast, wait until a critical need arises before they begin to communicate, making for a very high workload (Orasanu, 1990).

As we begin to think about the introduction of a computer-based intelligent flight assistant in the cockpit, the concept of crew awareness is further complicated because there may be critical information that this system takes account of that is not made available to the human crewmembers. The design of cockpit displays for such an assistant must consider the needs for crew awareness as well as for particular individual crewmembers.

The objective measurement of crew awareness is only beginning to be addressed. It will require the simultaneous and independent assessment of the awareness of each crewmember, using one or more of the methods we have discussed, and the correlation of the individual results. It may also require the scoring of the content of communication messages to identify the points in time at which information is shared.

SUMMARY AND CONCLUSIONS

Situation awareness, as a feature of aircrew performance, is receiving increased attention as new display technologies become available and higher levels of automation are being proposed for flight deck implementation. As a brief, comprehensive definition we quote Endsley,

"The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (Endsley, 1988).

Measuring situation awareness requires understanding the processes underlying the achievement of it as well as the kinds of informational "products" that result. These processes are characterized by a range of cognitive activities from automatic effortless processing to complex effortful inferential reasoning. Successful achievement of situation awareness may also be studied by examining the different methods used by experts to minimize the impact of excessive attention demands.

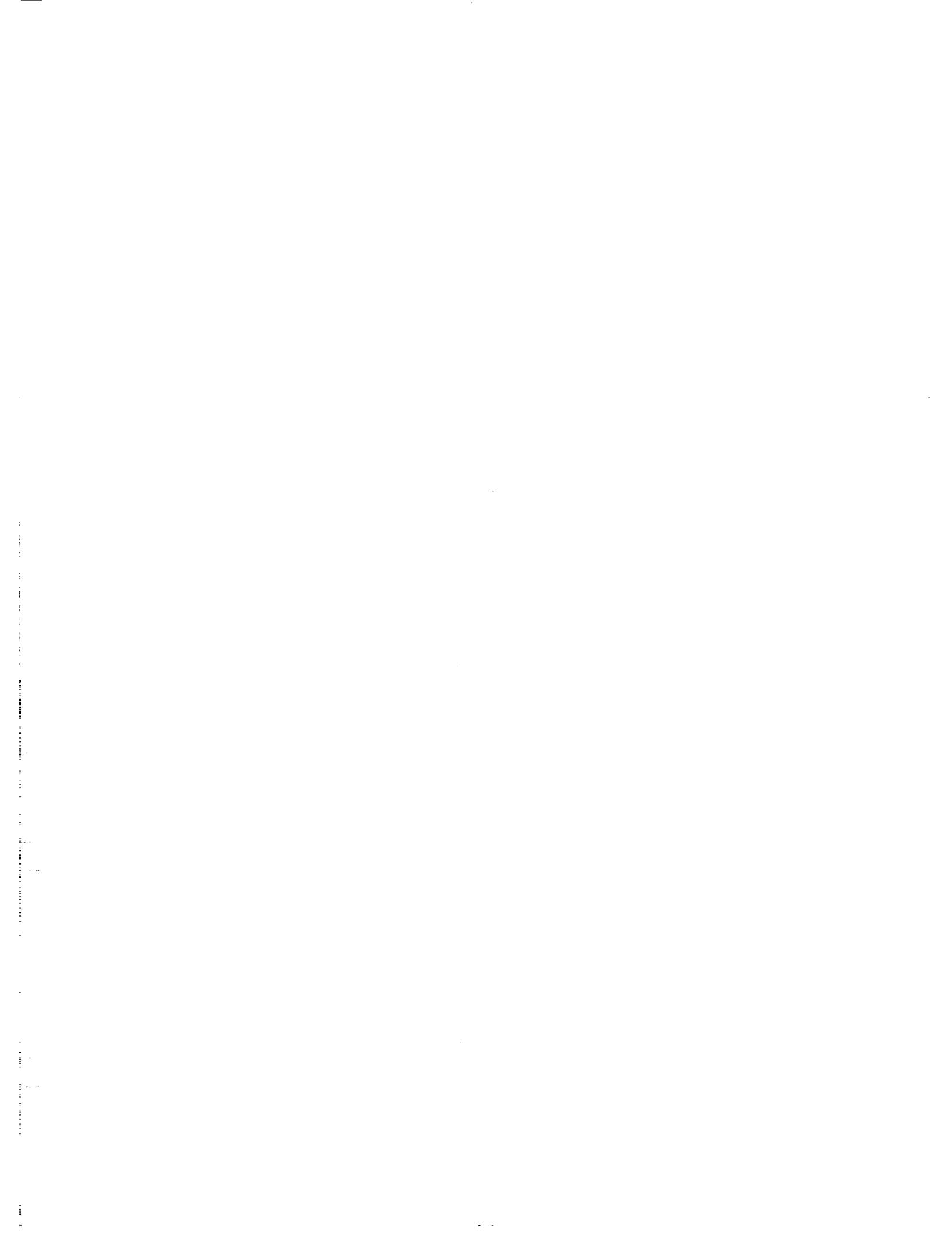
We emphasize three approaches to measurement strategy: (1) creative design of scenario contexts that emphasize the demand for the aircrew's attentional resources and the need to interpret subtle and sometimes misleading clues, (2) direct probes and experimental manipulations designed to tap the knowledge and strategies used by the aircrew during a scenario or brief vignette, and (3) inferences from a computer program that models the actions, knowledge states and information processing strategies of the crewmember.

The design of flight deck-related automation can have many subtle effects on the requirements for and ease of achieving situation awareness because, at once, it adds elements to the array of systems that must be monitored at the same time that it takes away information that previously was obtained by direct interaction with the aircraft. Therefore, improved design of human-centered automation can be facilitated by systematic objective measurement of the situation awareness associated with new designs.

REFERENCES

- Abbott, K. (1990). Robust Diagnosis in Physical Systems in Operation, Ph.D. Dissertation, Department of Computer Science, Rutgers University.
- Anderson, R. C., & Pitchard, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. Journal of Verbal Learning and Verbal Behavior, 17, 1-12.
- Baron, S., Zacharias, G., Muralidharan, R., and Landcraft, R. (1980). PROCRU: A model for analyzing flight crew procedures in approach to landing. Proceedings of the 8th IFAC World Congress, Tokyo, Japan.
- Bartlett, F. C. (1932). Remembering. Cambridge, England: Cambridge University Press.
- Beringer, D. B., & Hancock, P. A. (1989). Exploring situational awareness: A review and the effects of stress on rectilinear normalization. Proceedings of the 5th International Symposium on Aviation Psychology, Columbus, Ohio.
- Bower, G. H. (1982). Plans and goals in understanding episodes. In A. Flammer & W. Kintsch (Eds.), Discourse processing. New York: North-Holland.
- Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. Psychological Review, 64, 205-215.
- Chase, W.G., & Simon, H. A. (1973). Perception in chess. Cognitive Psychology, 4, 55-81.
- Corker, K.M., Cramer, N.L., & Henry, E. H. (1990). Methodology for evaluation of automation impacts on command and control systems. BBN Report 7242. Cambridge, MA: BBN Systems & Technologies.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. Proceedings of the Human Factors Society 32nd Annual Meeting.
- Endsley, M. R. (1989). A methodology for the objective measurement of situation awareness. Paper presented at the AGARD Symposium on Situation Awareness in Aerospace Operations, Copenhagen, Denmark.
- Fracker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. Proceedings of the Human Factors Society 32nd Annual Meeting, Santa Monica, CA.
- Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton-Mifflin.
- Hudlicka, E., Corker, K., Schudy, R., & Baron, S. (1989). Intelligent Aids for Flight Crew Tasks, BBN Report 6990. Cambridge, MA: BBN Systems & Technologies.
- Neisser, U. (1976). Cognition and reality: Principles and implications of cognitive psychology. San Francisco: W. H. Freeman.
- Norman, D. A. (1986). New views of information processing: Implications for intelligent decision support systems. In

- E. Hollnagel (Ed.), NATO ASI Series, Vol. F21. Berlin: Springer-Verlag.
- NTSB (1988). Aircraft accident report-- Northwest Airlines, McDonnell Douglas DC-9-82, N312RC, Detroit Metropolitan Wayne Co. Airport, Romulus MI, August 16, 1987. NTSB: Washington, DC.
- NTSB (1989). Aircraft accident report-- British Midland Airways, Boeing-737-400, G-QBME, East Midland Airport, January 8, 1989. NTSB: Washington, DC.
- Orasanu, J. M. (1990). Shared mental models and crew decision making. CSL Report 46. Princeton: Cognitive Science Laboratory, Princeton University.
- Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. Phil. Trans. R. Soc. of Lond. B, 327, 475-484.
- Regal, D. M., Rogers, W. H., & Boucek, G. P. (1988). Situational awareness in the commercial flight deck: Definition, measurement, and enhancement. Paper presented at the Aerospace Technology Conference and Exposition, Anaheim, California.
- Sanford, A. J., & Garrod, S.C. (1981). Understanding written language. New York: Wiley.
- Sarter, N. B., & Woods, D. D. (1991). Situation awareness. A critical but ill-defined phenomenon. International Journal of Aviation Psychology, 1, 45-57.
- Schank, R.C., Collins, G. C., Davis E., Johnson, P.N., Lytinen, S., & Reiser, B.J. (1982). What's the point? Cognitive Science, 6, 255-275.
- Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. Human Factors, 30, 539-566.
- Stockton, W. (1988). Trouble in the cockpit: The airlines tackle pilot error. The New York Times Magazine, March 27.
- Swets, J.A. (1984). Mathematical models of attention. In R. Parasuraman & R. Davies (Eds.), Varieties of attention. New York: Academic Press.
- Vicente, K. J. (1988). Adapting the memory recall paradigm to evaluate interfaces. Acta Psychologica, 69, 249-278.
- Vicente, K.J., & Rasmussen J. (1990). The ecology of human-machine systems II: Mediating "direct perception" in complex work domains. Ecological Psychology, 2, 207-249.
- Wellens, A. R. (1989). Effects of telecommunication media upon information sharing and team performance: Some theoretical and empirical observations. IEEE Aerospace and Electronic Systems, 4, 13-19.
- Wellens, A.. R. (1990). Assessing multi-person and person-machine distributed decision making using an extended psychological distancing model. Final Report No. AAMRL-TR-90-006. Ohio: Wright-Patterson Air Force Base.
- Whitaker, L. A., & Klein, G. A. (1988). Situation awareness in the virtual world: Situation assessment report. Proceedings of the 11th Symposium of Psychology in the Department of Defense.
- Woods, D. D., Wise, J. A., & Hanes, L. F. (1982). Evaluation of safety parameter display concepts, NP-2239. Palo Alto, CA: Electric Power Research Institute.





REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

| | | | |
|---|--|---|----------------------------|
| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED | |
| | July 1992 | Contractor Report | |
| 4. TITLE AND SUBTITLE | | 5. FUNDING NUMBERS | |
| A Principled Approach to the Measurement of Situation Awareness in Commercial Aviation | | C NAS1-18788 WU 505-64-13-21 | |
| 6. AUTHOR(S) | | | |
| Yvette J. Tenney Marilyn Jager Adams Richard W. Pew | | A.W.F. Huggins William H. Rogers | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) | | 8. PERFORMING ORGANIZATION REPORT NUMBER | |
| BBN Systems and Technologies 10 Moulton Street Cambridge, Massachusetts 02138 | | BBN Report No. 7451 | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER | |
| National Aeronautics and Space Administration Langley Research Center Hampton, VA 23665-5225 | | NASA CR-4451 | |
| 11. SUPPLEMENTARY NOTES | | | |
| Langley Technical Monitor: Kathy H. Abbott | | | |
| 12a. DISTRIBUTION/AVAILABILITY STATEMENT | | 12b. DISTRIBUTION CODE | |
| Unclassified - Unlimited Subject Category 53 | | | |
| 13. ABSTRACT (Maximum 200 words) The issue of how to support situation awareness among crews of modern commercial aircraft is becoming especially important with the introduction of automation in the form of sophisticated flight management computers and expert systems designed to assist the crew. In this paper, cognitive theories are discussed that have relevance for the definition and measurement of situation awareness. These theories suggest that comprehension of the flow of events is an active process that is limited by the modularity of attention and memory constraints, but can be enhanced by expert knowledge and strategies. Three implications of this perspective for assessing and improving situation awareness are considered: (1) Scenario variations are proposed that tax awareness by placing demands on attention; (2) Experimental tasks and probes are described for assessing the cognitive processes that underlie situation awareness; and (3) The use of computer-based human performance models to augment the measures of situation awareness derived from performance data is explored. Finally, two potential example applications of the proposed assessment techniques are described, one concerning spatial awareness using wide field of view displays and the other emphasizing fault management in aircraft systems. | | | |
| 14. SUBJECT TERMS | | 15. NUMBER OF PAGES | |
| situation awareness crew awareness pilot error | | automation and situation awareness perception attention | 32 |
| | | 16. PRICE CODE | |
| | | A03 | |
| 17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED | 18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED | 19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED | 20. LIMITATION OF ABSTRACT |